

UNITED STATES PATENT APPLICATION

FOR

METHOD AND APPARATUS FOR MULTIPLE CONTEXTS AND LAYER 3 VIRTUAL PRIVATE
NETWORKS

Inventors:

Ravi Chandra
Enke Chen
Jenny Yuan

Prepared by:

Blakely, Sokoloff, Taylor & Zafman LLP
12400 Wilshire Blvd., Suite 700
Los Angeles, California 90025
(408) 720-3800

EXPRESS MAIL CERTIFICATE OF MAILING

"Express Mail" mailing label number: EL 863955641 Date of Deposit: 11-17-01

I hereby certify that I am causing this paper or fee to be deposited with the United States Postal Service "Express Mail Post Office to Addressee" service on the date indicated above and that this paper or fee has been addressed to the Commissioner of Patents and Trademarks, Washington, D. C. 20241

Shenise Ramdeen

(Typed or printed name of person mailing paper or fee)

Shenise Ramdeen

(Signature of person mailing paper or fee)

11-17-01

(Date signed)

067

Field of the Invention

Background of the Invention

[0002] Virtual Private Networks (VPNs) extend an entity's (e.g., a corporation, Internet Service Provider (ISP), etc.) network backbone out to the Internet. The connectivity costs for VPNs are less than leasing a line, and fault tolerance is improved because of multiple pathways between sites. Instead of an entity purchasing, administrating and maintaining additional network elements (e.g. routers, hubs, switches, subscriber management systems, etc.), an entity can securely transmit traffic through the Internet with VPNs.

BRIEF DESCRIPTION OF THE DRAWINGS

[0003] The invention may best be understood by referring to the following description and accompanying drawings that are used to illustrate embodiments of the invention. In the drawings:

[0004] Figure 1 is a diagram of an exemplary network according to one embodiment of the invention.

[0005] Figure 2A is a diagram of an exemplary network element that supports multiple contexts and layer 3 VPNs according to one embodiment of the invention.

[0006] Figure 2B is a diagram of the exemplary network element 105 illustrating a relationship between network processes and the structures and tables illustrated in Figure 2A according to one embodiment of the invention.

[0007] Figure 2C is a diagram of the layer 3 VPN EGP table 215 as a tree according to one embodiment of the invention.

[0008] Figure 2D is a diagram of the layer 3 VPN EGP table 215 as a tree with RD substructures according to one embodiment of the invention.

[0009] Figure 2E is a diagram illustrating an alternative embodiment of the layer 3 VPN EGP table 215 as a hash table according to one embodiment of the invention.

[0010] Figure 3A is a flow chart for processing an update message received from customer with RD substructures equipment according to one embodiment of the invention.

[0011] Figure 3B is a flow chart continuing from the flow chart of Figure 3A according to one embodiment of the invention.

[0012] Figure 4A is a flow chart for processing an update message received from customer equipment according to one embodiment of the invention.

[0013] Figure 4B is a flow chart continuing from the flow chart illustrated in Figure 4A according to one embodiment of the invention.

DETAILED DESCRIPTION OF THE INVENTION

[0018] In the following description, numerous specific details are set forth to provide a thorough understanding of the invention. However, it is understood that the invention may be practiced without these specific details. In other instances, well-known circuits, structures, standards, and techniques have not been shown in detail in order not to obscure the invention.

[0019] Figure 1 is a diagram of an exemplary network according to one embodiment of the invention. In Figure 1, a virtual private network (VPN) A includes a VPN A customer equipment 101A and a VPN A customer equipment 103A. A VPN B includes a VPN B customer equipment 101B and a VPN B customer equipment 103B. The VPN A customer equipment 101A and the VPN B customer equipment 103A are located within an autonomous system (AS) 102 in Figure 1, but could be located in different ASs. The VPN A customer equipment 101A and the VPN B customer equipment 103A are coupled with a network element 105. The network element 105 is also coupled with a non-VPN customer equipment 106 that is located within an AS 104. The network element 105 is located in an AS that is a backbone 111.

[0020] A network element 115 is also located within the backbone 111. The network element 115 is coupled with the VPN B customer equipment 103B. A network element 114, also located in the backbone 111, is coupled with the VPN A customer equipment 101B. The VPN B customer equipment 103B and the VPN A customer equipment 101B are respectively located within an AS 106 and an AS 108.

[0021] The network element 105 includes a context 107A and VPN contexts 109A – 109B. The term context is used herein to refer to a set of information and/or collection of data structures for a customer of a network provider's network element. In the exemplary network illustrated in Figure 1, the VPN A customer equipment 101A, the VPN B customer equipment 103A, and the non-VPN customer equipment 106 gain access to the backbone 111

The figure consists of ten vertical panels, each containing a black-and-white micrograph of an embryo at a specific developmental stage. The panels are numbered 1 through 10 from top to bottom. Panel 1 shows a very early stage, likely a zygote or cleavage stage. Panels 2-4 show increasing complexity as the embryo develops. Panel 5 shows a more defined body plan. Panel 6 shows a distinct head and tail region. Panel 7 shows further differentiation of internal structures. Panel 8 shows a well-defined larval form. Panel 9 shows a slightly larger larva. Panel 10 shows a fully developed larva, possibly ready to hatch.

[0025] Implementing VPNs at layer 3 provides numerous advantages. Layer 3 VPNs provide vast scalability and avoids administrative costs of tunnel configurations for layer 2 VPNs. Combining layer 3 VPNs with multiple contexts provides a powerful tool for servicing various customers that may include VPN customers and non-VPN customers. A network element supporting multiple contexts and layer 3 VPNs can support a large number of VPN customer and non-VPN customers and provide each customer access to their context.

[0026] Figure 2A is a diagram of an exemplary network element that supports multiple contexts and layer 3 VPNs according to one embodiment of the invention. Figure 2 illustrates the network element 105 illustrated in Figure 1. Figure 2, also illustrates the context 107A and the VPN context 109A and 109B illustrated in Figure 1. The context 107A includes a peer data structure 201A, a configuration data structure 203A, a routing table 205A, an interior gateway protocol (IGP) table 207A, and an exterior gateway protocol (EGP) table 209A. The term table is a conventional term and not meant to be limiting upon the type of data structure. A table can be implemented with different data structures such as a hash table, search tree, combinations of data structures, etc.

[0027] The peer structure 201A includes data describing a peer. A peer is another network element that exchanges routing information with the network element 105. A peer

can either be an internal peer or an external peer. An internal peer is a network element directly connected to the network element 105 that is located within the same autonomous system as the network element 105. An autonomous system is a collection of network elements under a single administrative authority that routes traffic in accordance with a common IGP (e.g., OSPF, IS-IS, RIP, etc.). An external peer of the network element 105 is a network element that is directly connected to the network element 105, but is located in a different autonomous system. The peer structure 201A may include a peer's network address and configuration information providing guidelines for communications between the peer and the network element 105.

[0028] The configuration structure 203A includes configurations specifically for the context 107A. The configurations in the configuration structure 203A may be entered by a network administrator of the network element 105 or the administrator for the owner of the context 107A. The routing table 205A is a routing table specifically for the context 107A. The routing table 205A is used to process traffic transmitted and received for the context 107A. The IGP table 207A is a forwarding table specifically for the context 107A. The IGP table 207A may be an OSPF forwarding table, a RIP forwarding table, an ISIS forwarding table, etc. The exterior gateway protocol forwarding table 209A is a forwarding table with routes specifically for the context 107A maintained by an EGP process (e.g., BGP, etc.). As previously stated, the owner of the context 107A can be given permission by the owner of the network element 105 to view and manipulate information for the context 107A.

[0029] Since the number of entries in a table for a non-VPN customer is typically large, separation of tables for non-VPN customers provides for efficient maintenance of the non-VPN tables. Since the number of entries in a VPN table are typically fewer than the number of entries in VPN tables, sharing a single EGP table for multiple VPNs provides for efficient walking of the shared VPN tables as describe in more detail later herein.

[0030] Various embodiments of the invention may implement the EGP tables, IGP tables, and routing tables differently. For example, one embodiment may maintain a single EGP table for non-VPN contexts and a single EGP table for layer 3 VPN contexts. Another embodiment may maintain a single IGP table for VPN contexts.

[0031] The VPN context 109A includes a peer structure 201I, a configuration structure 203I, a routing table 205I, and an IGP table 207I. Similarly, the VPN context 109B includes a peer structure 201J, a configuration structure 203J, a routing table structure 205J, and an IGP table 207J. The information included in the structures 201I and 203I and the tables 205I and 207I are similar to the structures 201A and 203A and the tables 205A and 207A, but are specific to the VPN context 109A. Likewise, the structures 201J and 203J and the tables 205J and 207J are specific to the VPN context 109B. As previously stated, the owners of the VPN context 109A and 109B can respectively access and configure the structures and tables in the VPN contexts 109A and 109B. Unlike the context 107A, the VPN contexts 109A and 109B do not have specific EGP tables. Instead, the VPN contexts 109A and 109B share a layer 3 VPN EGP table 215. The layer 3 VPN EGP table 215 includes EGP forwarding information for all VPN contexts in the network element 105.

[0032] In addition to the context 107A and the VPN contexts 109A and 109B, in one embodiment the network element 105 includes a local context 211. A peer structure 201X, a configuration structure 203X, a routing table 205X, an IGP table 207X, and an EGP table 209X contain information for the local context 211. The information included in the structure 201X and the tables 205X, 207X, and 209X is similar to the information included in the structure 201A and the tables 205A, 207A, and 209A, but is specific to the local context 211. The configuration structure 203X includes global configuration information (e.g., system time). Unlike the context 107A and the VPN contexts 109A and 109B, in one embodiment the owner of the local context 211, who is the owner of the network element

105, has access to many of the structures and tables of the contexts on the network element 105.

[0033] Figure 2B is a diagram of the exemplary network element 105 illustrating a relationship between network processes and the structures and tables illustrated in Figure 2A according to one embodiment of the invention. In Figure 2B, a configuration interface (CI) 221 feeds configuration information into a configuration process 223. The CI may be a command line interface or a graphical user interface for entering configurations and commands for the network element 105. The configuration process 223 stores the configurations into one of the configuration structures 203X, 203A, 203I, and 203J. The configuration structures 203X, 203A, 203I, and 203J are illustrated in Figure 2B as referencing the peer structures 201X, 201A, 201I, and 201J. The relationships are illustrated as follows : the peer structure 201J and the configuration structure 203J; the peer structure 201I and the configuration structure 203I; the peer structure 201A and the configuration structure 203A; and the peer structure 201X and the configuration structure 203X.

[0034] The peer structures may be implemented in a variety of ways. In one embodiment, a single peer structure has entries for all peers configured on a given network element. In such an embodiment, each entry identifies a peer and its corresponding context. In another embodiment, each peer configured on the given network element is allocated a single peer structure. In another embodiment, a single peer structure is allocated for each context configured on the network element.

[0035] The configuration process 223 also sends configurations to an IGP process 225, a RIB process 229, and an EGP process 227. The IGP process 225 represents one or more IGP processes that may include the following protocols: RIP, OSPF, IS-IS, etc. The EGP process 227 also represents one or more processes that may include BGP and other exterior gateway protocols. The IGP process 225 maintains the IGP tables 207X, 207A, 207I, and 207J. The

RIB process 229 maintains the routing tables 205X, 205A, 205I, and 205J. The EGP process 227 maintains the EGP table 209X, 209A, the layer 3 VPN EGP table 215, and a mapping structure 226.

[0036] In an embodiment that provides for a single non-VPN EGP table and a single VPN EGP table, an identifier (e.g., context ID) is included in each entry of the single non-VPN EGP table. In another embodiment that provides for single non-VPN EGP table and a single VPN EGP table, each entry of the non-VPN EGP table indicates an identifier (e.g., context ID) and references a set of prefixes corresponding to the identifier.

[0037] The mapping structure 226 is a data structure that includes VPN context IDs and their corresponding RD. The mapping structure 226 may be implemented as a tree, hash table, linked list, etc. When the IGP process 225 downloads updates of one of the IGP tables, such as the IGP table 207X, the IGP process 225 passes the context ID indicated in the IGP table 207X to the RIB process 229. The EGP process 227 does the same for non-VPN EGP tables. However, the layer 3 VPN EGP table 215 does not correspond to a single context. Therefore, the EGP process 227 uses the mapping structure 226 to communicate with other processes in the network element 105 for communications related to VPN contexts. Communications with other processes may include passing updates to the RIB process 229, performing requests from the configuration process 223, etc.

[0038] Sharing a single EGP table for all of the VPN contexts of a network element improves the efficiency of update message generation for the network element supporting multiple contexts and layer 3 VPNs. Since the number of entries for an individual VPN customer is relatively small, combining entries for all layer 3 VPN customers into a single EGP table enables an EGP process to efficiently walk the combined data structure instead of numerous small EGP tables. In addition, a mapping structure enables efficient maintenance of the combined EGP table without extensive memory consumption.

[0039] Figure 2C is a diagram of the layer 3 VPN EGP table 215 as a tree according to one embodiment of the invention. In Figure 2C, the layer 3 VPN EGP table is implemented as a tree. Each of the nodes 241, 243, 245, 247, and 249 are labeled with the value RD:prefix. The mapping structure 226 is also illustrated in Figure 2C. The EGP process 227 uses the mapping structure 226 as previously discussed to send and receive communication regarding the layer 3 VPN EGP table 215. If the EGP process 227 receives a communication affecting the layer 3 VPN EGP table 215, then the EGP process 227 maps the VPN-context ID indicated in the communication to a RD with the mapping structure 216. Maintaining the mapping structure 226 avoids including a VPN context ID in every node.

[0040] Figure 2D is a diagram of the layer 3 VPN EGP table 215 as a tree with RD substructures according to one embodiment of the invention. Unlike the layer 3 VPN EGP table 215 illustrated in Figure 2C, the right branch of the root node 241 references a node 251. Instead of RD:prefix, the node 251 is labeled with the route distinguisher. In addition to typical left and right branches of a tree, the node 251 references an RD substructure 257. The RD substructure 257 is a data structure that indicates prefixes corresponding to the RD identified in the node 251. A node 253 is similar to the node 251 in that it is only identified by the value RD instead of the value RD:prefix. The node 253 references an RD substructure 259. The RD substructure 259 includes prefixes corresponding to the RD identified in the node 253. The embodiment of the layer 3 VPN EGP table 215 further improves memory use. The embodiment illustrated in Figure 2D reduces the number of nodes in the layer 3 VPN EGP table 215 with the route distinguisher value (currently defined as an 8 byte value), and group prefixes by route distinguisher. In addition, the EGP process 227 may group updates to be transmitted by RD as it encounters each node of the layer 3 VPN EGP table 215.

[0041] Figure 2E is a diagram illustrating an alternative embodiment of the layer 3 VPN EGP table 215 as a hash table according to one embodiment of the invention. Each entry of

the hash table 270 indicates a route distinguisher. Each entry of the hash table 260 references an RD substructure 271 or 273. As with the layer 3 VPN EGP table illustrated in Figure 2D, the EGP process 227 may group updates by mapped VPN context ID as it walks through the hash table 260.

[0042] As illustrated in Figures 2C-2E, various embodiments of the invention may implement the layer 3 VPN EGP table 215 with different data structures and/or combinations of data structures.

[0043] Figure 3A is a flow chart for processing an update message received from customer with RD substructures equipment according to one embodiment of the invention. At block 301, the EGP process 227 selects a peer that is a customer equipment. At block 303, the EGP process 227 receives an update message from the selected peer. At block 305, the EGP process 227 selects a prefix from the received update message. At block 307, the EGP process 227 determines if the context associated with the selected peer is a VPN context. If the associated context is not a VPN context, then at block 309 the EGP process 227 determines if the selected prefix is in the EGP table of the context associated with the selected peer. If the EGP process 227 determines that the selected prefix is within the associated context's EGP table, then control flows to block 313. If the EGP process 227 determines that the selected prefix is not within the associated context's EGP table, then control flows to block 315.

[0044] If at block 309 the EGP process 227 determines that the context associated with the selected peer is a VPN context, then at block 317 the EGP process 227 maps the VPN context ID of the associated context to a route distinguisher (RD). At block 318 the EGP process 227 locates the RD in the layer 3 VPN EGP table 215. At block 319, the EGP process 227 determines if the selected prefix is in the located RD sub-structure. If the EGP process 227 determines that the selected prefix is within the RD sub-structure, then control

flows to block 323. If the EGP process 227 determines that the selected prefix is not within the located RD sub-structure, then control flows to block 321.

[0045] Figure 3B is a flow chart continuing from the flow chart of Figure 3A according to one embodiment of the invention. At block 313, the EGP process 227 inserts the selected prefix into the selected context EGP table. At block 315, the EGP process 227 updates an entry matching the selected prefix of the associated context's EGP table in accordance with the update message. At block 321, the EGP process 227 inserts a selected prefix into the located RD substructure. At block 323, the EGP process 227 updates an entry matching the selected prefix within the located RD substructure in accordance with the update message.

[0046] Figure 4A is a flow chart for processing an update message received from customer equipment according to one embodiment of the invention. At block 401, the EGP process 227 selects a peer that is a customer equipment. At block 403, the EGP process 227 receives an update message from the selected peer. At block 405, the EGP process 227 selects a prefix from the received update message. At block 407, the EGP process 227 determines if the context associated with the selected peer is a VPN context. If the associated context is not a VPN context, then at block 409 the EGP process 227 determines if the selected prefix is in the EGP table of the context associated with the selected peer. If the EGP process 227 determines that the selected prefix is within the associated context's EGP table, then control flows to block 415. If the EGP process 227 determines that the selected prefix is not within the associated context's EGP table, then control flows to block 413.

[0047] If at block 407, the EGP process 227 determines that the selected context is a VPN context, then at block 417 the EGP process 227 maps the VPN context ID of the associated context to a (RD). At block 419 the EGP process 227 determines if the RD:selected prefix is within the layer 3 VPN EGP table 215. If the RD:selected prefix is within the later 3 VPN

[0048] Figure 4B is a flow chart continuing from the flow chart illustrated in Figure 4A according to one embodiment of the invention. At block 421, the EGP process 227 inserts RD:selected prefix into the layer 3 VPN EGP table 215. At block 423, the network element updates an entry matching the selected prefix in the layer 3 VPN EGP table in accordance with the update message. At block 413, the EGP process 227 inserts the selected prefix into the associated context's EGP table. At block 415, the EGP process 227 updates an entry matching a selected prefix in the associated context's EGP table in accordance with the update message.

[0050] If at block 507, the EGP process 227 determines that the selected peer is not configured for the VPN address family, then at block 511 the EGP process 227 determines if the selected prefix matches an entry in the layer 3 VPN EGP table 215. If the selected prefix

does not match an entry in the VPN EGP table 215, then at block 515 the EGP process 227 inserts the select prefix into the layer 3 VPN EGP table 215. If the selected prefix matches an entry in the layer 3 VPN EGP table 215, then at block 513 the EGP process 227 updates the matching entry in accordance with the update message.

[0051] Figure 6 is a flow chart for generating an update message for customer equipment according to one embodiment of the invention. At block 601, the EGP process 227 selects a peer. At block 603, the EGP process 227 determines if the context associated with the selected peer is a VPN context. If the associated context is not a VPN context, then at block 607, the EGP process 227 creates an update message(s) for the associated context's EGP table. If the EGP process 227 determines at block 603 that the associated context is a VPN context, then at block 609 the EGP process 227 maps the VPN context ID of the associated VPN context to an RD. At block 613, the EGP process 227 locates the RD in the VPN EGP table 215. At block 615, the EGP process 227 walks through the VPN EGP table and select entries matching the mapped RD. At block 617, the EGP process 227 creates an update message(s) with prefixes of selected entries without the RD.

[0052] Figure 7 is a flow chart for generating an update message for a provider equipment according to one embodiment of the invention. At block 701, the EGP process 227 selects a peer. At block 703, the EGP process 227 determines if the selected peer is configured for the VPN address family. If the selected peer is not configured for the VPN address family, then the EGP process 227 creates an update message(s) with the associated context's EGP table at block 705. If the selected peer is not configured for the VPN address family, then the EGP process 227 walks through the layer 3 VPN EGP table 215 to create update message(s) at block 707.

[0053] While the flow diagrams in the Figures show a particular order of operations performed by certain embodiments of the invention, it should be understood that such order is

exemplary (e.g., alternative embodiments may perform certain of the operations in a different order, combine certain of the operations, perform certain of the operations in parallel, etc.).

[0054] Figure 8 is a block diagram illustrating the exemplary network element 105 according to one embodiment of the invention. In Figure 8, a control card 801 hosts the local context 211, the VPN context 109A, the VPN context 109B, the context 107A, and the Layer 3 VPN EGP table 215, the configuration process 223, the IGP process 225, the EGP process 227, and the RIB process 229. In various embodiments of the invention, the local context 211, the VPN context 109A, the VPN context 109B, the context 107, and the Layer 3 VPN EGP table 215, the configuration process 223, the IGP process 225, the EGP process 227, and the RIB process 229 may be hosted on a co-processor, an ASIC, etc. In alternative embodiments, the structures and tables of the context 107A and VPN contexts 109A and 109B may be stored on a storage device and/or memory coupled with the control card 801.

[0055] The control card 801 is coupled with a transmission medium cloud 803 (e.g., a system bus, a mesh, a system bus and a mesh, etc.). The transmission medium cloud 803 is coupled with line cards 807A – 807D. The line cards 807A – 807D are coupled to physical interfaces 809A – 809D respectively. The network element 105 receives updates messages and transmits update messages via the physical interfaces 805A – 805D.

[0056] The control card 801 and the line cards 807A – 807D illustrated in Figure 8 includes memories, processors, and/or ASICs. Such memories include a machine-readable medium on which is stored a set of instructions (i.e., software) embodying any one, or all, of the methodologies described herein. Software can reside, completely or at least partially, within this memory and/or within the processor and/or ASICs. For the purpose of this specification, the term "machine-readable medium" shall be taken to include any mechanism that provides (i.e., stores and/or transmits) information in a form readable by a machine (e.g., a computer). For example, a machine-readable medium includes read only memory

("ROM"), random access memory ("RAM"), magnetic disk storage media, optical storage media, flash memory devices, electrical, optical, acoustical, or other form of propagated signals (e.g., carrier waves, infrared signals, digital signals, etc.), etc.

[0057] A network element supporting multiple contexts and layer 3 VPNs provides the owner of the network element the ability to support VPNs for multiple customers at a cost lower than the cost of layer 2 VPNs and more efficiently than without multiple contexts. The owner can provide customers access to their own context. Providing access to contexts can be offered as a service and shifts some administrative tasks from the network element owner to its customers. Hence, the network element owner can provide layer 3 VPNs to multiple VPN customers (e.g. corporations, service providers, education institutions, etc.) and backbone access to non-VPN customers. The network element owner can serve this variety of customers from a single network element while offering administrative access to the network element through multiple contexts.

[0058] The method and apparatus of the invention can be practiced with modification and alteration within the spirit and scope of the appended claims. The description is thus to be regarded as illustrative instead of limiting on the invention.